MONO- AND DIACYLGLYCEROL ACYLTRANSFERASES AND METHODS OF USE THEREOF

CROSS REFERENCE

This application is a continuation-in-part of U.S. Patent Application Serial No. 09/794,715, filed February 26, 2001, which application claims priority to the filing date of the U.S. Provisional Patent Application Serial No. 60/271,307, filed February 23, 2001, the disclosures of which are herein incorporated by reference in their entirety.

GOVERNMENT RIGHTS

The United States Government may have certain rights in this application pursuant to Grant No. DK56084 from the National Institutes of Health.

FIELD OF THE INVENTION

The field of the invention is enzymes, particularly acyltransferases.

BACKGROUND OF THE INVENTION

Diacylglycerol O-Acyltransferase (EC 2.3.1.20), also known as diglyceride acyltransferase or DGAT, is a critical enzyme in triacylglycerol synthesis. Triacylglycerols are quantitatively the most important storage form of energy for eukaryotic cells. DGAT catalyzes the rate-limiting and terminal step in triacylglycerol synthesis using diacylglycerol and fatty acyl CoA as substrates. As such, DGAT plays a fundamental role in the metabolism of cellular diacylglycerol and is important in higher eukaryotes for intestinal fat absorption, lipoprotein assembly, fat storage in adipocytes, milk production and possibly egg production

and sperm maturation.

Diacylglycerol is the precursor of such important lipids as triacylglycerol and phospholipids, which store energy and form cellular membranes. In eukaryotes, two major pathways for synthesizing diacylglycerol exist: the glycerol phosphate pathway and the monoacylglycerol pathway. Both pathways generate diacylglycerol that can be used as a substrate by acyl CoA:diacylglycerol acyltransferase (DGAT) for triacylglycerol synthesis. In the glycerol phosphate pathway, which functions in most cells, diacylglycerol is derived by the dephosphorylation of phosphatidic acid produced by sequential acylations of glycerol phosphate. In the monoacylglycerol pathway, which has been reported predominantly in the intestine, diacylglycerol is formed directly from monoacylglycerol and fatty acyl CoA in a reaction catalyzed by monoacylglycerol acyltransferase (MGAT) (E.C. 2.3.1.22).

MGAT is best known for its role in fat absorption in the intestine, where the fatty acids and *sn*-2-monoacylglycerol generated from the digestion of dietary fat (mainly triacylglycerol) are resynthesized into triacylglycerol in enterocytes for chylomicron synthesis and secretion. MGAT catalyzes the first step of this process, in which fatty acyl CoA, formed from fatty acids and CoA, and *sn*-2-monoacylglycerol are covalently joined. Because the monoacylglycerol pathway predominates in intestinal triacylglycerol synthesis, MGAT may be a pharmaceutical target for modulating fat absorption.

MGAT activity is also found at high levels in liver of suckling rats and in white adipose tissue of migrating sparrows, where triacylglycerols are actively hydrolyzed to provide fatty acids for energy. MGAT preferentially acylates monoacylglycerols that contain a polyunsaturated fatty acyl moeity at the *sn*-2 position. Thus, MGAT may preserve essential fatty acids, all of which are polyunsaturated, by resynthesizing them into triacylglycerols. This function may be relevant in mammalian white adipose tissue, which possesses significant levels of MGAT activity. In addition, MGAT may also play a role in signaling, since its product, diacylglycerol, and one of its substrates, 2-arachidonoylglycerol, are signaling molecules.

Like many enzymes that participate in neutral lipid synthesis, MGAT has proven difficult to purify to homogeneity, and an MGAT gene has not been identified. Several partial purifications of MGAT enzymes have been reported, and a 43-kDa MGAT enzyme was

purified recently from peanut cotyledons. Difficulties in the purification of MGAT may reflect its hydrophobicity or its involvement in an enzyme complex.

Because of its central role in a variety of different processes, there is much interest in the identification of polynucleotides encoding proteins having DGAT and MGAT activity, as well as the proteins encoded thereby.

Relevant Literature

Of particular interest are: U.S. Patent No. 6,100,077; and PCT Published Application Nos. WO 98/55631; WO 99/67268; WO 00/01713; WO 99/67403; WO 00/32793; WO 00/32756; WO 00/36114; WO 00/60095; WO 00/66749.

Also of interest are: Smith et al., Nat.Genet. 2000 (25), 87-90).Cases et al., "Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis," Proc. Natl. Acad. Sci. USA (October 1998) 95:13018-13023; Oelkers et al., "Characterization of Two Human Genes Encoding Acyl Coenzyme A: Cholesterol Acyltransferase-Related Enzymes," J. Biol. Chem. (October 9, 1998) 273:26765-71; and Cases et al. (2001) *J. Biol. Chem.* 276:38870-38876.

References describing the role DGAT plays in various biological processes include: Bell & Coleman, "Enzymes of Glycerolipid Synthesis in Eukaryotes," Annu. Rev. Biochem. (1980) 49: 459-487; Lehner & Kuksis, "Biosynthesis of Triacylglycerols," Prog. Lipid Res. (1996) 35: 169-201; Brindley, Biochemistry of Lipids, Lipoproteins and Membranes (eds. Vance & Vance)(Elsevier, Amsterdam)(1991) pp171-203; Haagsman & Van Golde, "Synthesis and Secretion of Very Low Density Lipoproteins by Isolated Rat Hepatocytes in Suspension: Role of Diacylglycerol Acyltransferase," Arch. Biochem. Biophys. (1981) 208:395-402; Coleman & Bell, "Triacylglycerol Synthesis in Isolated Fat Cells. Studies on the Microsomal Diacylglycerol Acyltransferase Activity Using Ethanol-Dispersed Diacylglycerols," J. Biol. Chem. (1976) 251:4537-4543.

References discussing MGAT activity and purification include: Coleman and Haynes (1984) *J. Biol. Chem.* 259:8934-8938; Mostafa et al. (1994) *Lipids* 29:785-791; Xia et al. 1993) *Am. J. Physiol.* 265:R414-R419; Jamdar et al. (1992) *Arch. Biochem. Biophys.* 296:419-425; Manganaro et al. (1985) *Can. J. Biochem. Cell Biol.* 63:341-347; Bhat et al. (1993) *Arch. Biochem. Biophys.* 300:663-669; Tumaney et al. (2001) *J. Biol. Chem.*

276:10847-10852; Lehner and Kuksis (1995) *J. Biol. Chem.* 270:13630-13636.

SUMMARY OF THE INVENTION

Nucleic acid compositions encoding polypeptide products with diglyceride acyltransferase and/or monoacylglycerol acyltransferase activity, as well as the polypeptide products encoded thereby, i.e., mammalian DGAT2α and MGAT1 polypeptide products, and methods for producing the same, are provided. Also provided are: methods and compositions for modulating DGAT2α and MGAT1 activity; DGAT2α and MGAT1 transgenic cells, animals and plants, as well as methods for their preparation; and methods for making diglyceride, diglyceride compositions, triglycerides and triglyceride compositions, as well as the compositions produced by these methods. The subject methods and compositions find use in a variety of different applications, including research, medicine, agriculture and industry applications.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 provides a graphical representation of the results obtained from a pulse assay that demonstrates the existence of mouse DGAT2α.

Fig. 2 provides a hydrophobicity plot of mouse DGAT2α.

Figs 3A to 3C provide graphical results of various mouse DGAT2α activity assays.

Fig. 4 provides the expression profile for mouse DGAT2α.

Fig. 5 provides the results of an assay showing that mouse DGAT2α expression increases during 3T3-L1adipocyte differentiation.

Figs 6A and 6B provide the amino acid and nucleic acid sequences of mouse DGAT2α.

Figs. 7A and 7B provide the amino acid and nucleic acid sequences of human DGAT2α.

Figures 8A-D provide the amino acid and nucleic acid sequences of various mouse and human DGAT2α homologs.

Figure 9 depicts schematically two major pathways for synthesizing diacylglycerol (DAG).

Figure 10A depicts a comparison of the amino acid sequences of mouse MGAT1 and mouse DGAT2. Figure 10B depicts a hydrophobicity plot of mouse MGAT1.

Figure 11 depicts graphs showing MGAT1 activity using oleoyl CoA or 2-monooleoylglycerol.

Figure 12 depicts graphs showing that MGAT1 has activity toward all stereoisomers of monoacyl glycerol.

Figure 13 depicts expression of MGAT1 protein in COS-7 cells.

Figure 14 depicts the tissue distribution of MGAT1 mRNA expression.

DETAILED DESCRIPTION OF THE INVENTION

Nucleic acid compositions encoding polypeptide products with diglyceride and/or monoacylglycerol acyltransferase activity, as well as the polypeptide products encoded thereby, i.e., mammalian DGAT2α and MGAT1, and methods for producing the same, are provided. In many embodiments, the subject nucleic acids encode enzymes that exhibit monoacylglycerol acyltransferase activity, diacylglycerol acyltransferase activity, or both mono- and diacyltransferase activity. For example, DGAT2α polypeptides exhibit diglyceride acyltransferase activity (also referred to herein as "DGAT2" polypeptides); and mammalian MGAT1 polypeptides (also referred to herein as "DC2" polypeptides) exhibit monoacylglycerol acyltransferase activity, and in some embodiments also exhibit diacylglycerol acyltransferase activity.

Also provided are: methods and compositions for modulating DGAT2α and MGAT1 activity, e.g. in the treatment of disease conditions associated with DGAT2α and/or MGAT1 activity, including obesity; MGAT1 and DGAT2α transgenic cells, animals, plants and fungi, and methods for their preparation, e.g. for use in research, food production, industrial feedstock production, etc.; and methods for making diglycerides, diglyceride compositions, triglycerides, and triglyceride compositions, e.g. oils. The methods and compositions of the subject invention find use in a variety of different applications and fields, including research, medicine, agriculture and industry.

Before the subject invention is further described, it is to be understood that the invention is not limited to the particular embodiments of the invention described below, as variations of the particular embodiments may be made and still fall within the scope of the appended claims. It is also to be understood that the terminology employed is for the purpose of describing particular embodiments, and is not intended to be limiting. Instead, the scope of the present invention will be established by the appended claims.

In this specification and the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention belongs.

NUCLEIC ACID COMPOSITIONS

Nucleic acid compositions encoding polypeptide products, as well as fragments thereof, having mono- and/or diglyceride acetyltransferase activity are provided. In many embodiments, the subject nucleic acids encode enzymes that exhibit monoacylglycerol acyltransferase activity, diacylglycerol acyltransferase activity, or both mono- and diacyltransferase activity. Specifically, nucleic acid compositions encoding mammalian, e.g., human, mouse, etc., DGAT2α polypeptides having diglyceride acyltransferase activity (also referred to herein as "DGAT2" polypeptides), and mammalian MGAT1 polypeptides exhibiting monoacylglycerol acyltransferase activity (also referred to herein as "DC2" polypeptides), are provided.

By nucleic acid composition is meant a composition comprising a sequence of DNA having an open reading frame that encodes a DGAT2α or an MGAT1 polypeptide, i.e. a gene or genomic region encoding a polypeptide having mono- and/or diglyceride acyltransferase activity, and is capable, under appropriate conditions, of being expressed as a DGAT2α or an MGAT1 polypeptide.

Also encompassed in this term are nucleic acids that are homologous or substantially

similar or identical to the nucleic acids encoding DGAT2α polypeptides, or MGAT1 polypeptides. Thus, the subject invention provides nucleic acids encoding mammalian DGAT2α, such as nucleic acids encoding human DGAT2α and homologs thereof and mouse DGAT2α and homologs thereof. The subject invention provides nucleic acids encoding mammalian MGAT1, such as nucleic acids encoding mouse MGAT1 (also referred to herein as "DC2"), and homologs thereof.

The coding sequence of the human DGAT2α genomic sequence, i.e. the human cDNA encoding the human DGAT2α enzyme, includes or comprises a nucleic acid sequence substantially the same as or identical to that identified as SEQ ID NO:01 or SEQ ID NO:18, *infra*. The coding sequence of the mouse DGAT2α genomic sequence, i.e., the mouse cDNA encoding the mouse DGAT2α enzyme, includes or comprises a nucleic acid substantially the same as or identical to the sequence identified as SEQ ID NO:03, *infra*. The coding sequence of the mouse MGAT1 genomic sequence, i.e. the mouse cDNA encoding the mouse MGAT1 enzyme, includes or comprises a nucleic acid sequence substantially the same as or identical to that identified as SEQ ID NO:05, *infra*.

The source of homologous nucleic acids to those specifically listed above may be any species, including both animal and plant species, e.g., primate species, particularly human; rodents, such as rats and mice, canines, felines, bovines, ovines, equines, yeast, nematodes, etc. Between mammalian species, *e.g.*, human and mouse, homologs have substantial sequence similarity, *e.g.* at least 75% sequence identity, usually at least 90%, more usually at least 95% between nucleotide sequences. Sequence similarity is calculated based on a reference sequence, which may be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, *etc.* A reference sequence will usually be at least about 18 nt long, more usually at least about 30 nt long, and may extend to the complete sequence that is being compared. Algorithms for sequence analysis are known in the art, such as BLAST, described in Altschul *et al.* (1990), *J. Mol. Biol.* **215**:403-10. Unless specified otherwise, all sequence identity values provided herein are determined using GCG (Genetics Computer Group, Wisconsin Package, Standard Settings, gap creation penalty 3.0, gap extension penalty 0.1). The sequences provided herein are essential for recognizing DGAT2α - related and

homologous polynucleotides in database searches. Specific DGAT2α homologues of interest are provide in Fig. 8, i.e., SEQ ID NOs. 05, 07, 09, 11, 13 and 15.

Also provided are nucleic acids that hybridize to the above-described specific nucleic acids, e.g., those nucleic acids having a sequence of SEQ ID NO:01, 03, 05, 07, 09, 11, 13, 15, or 18, or a coding sequence of any one of the foregoing sequences, under stringent conditions. An example of stringent hybridization conditions is hybridization at 50°C or higher and 0.1×SSC (15 mM sodium chloride/1.5 mM sodium citrate). Another example of stringent hybridization conditions is overnight incubation at 42°C in a solution: 50% formamide, 5 × SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5 × Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1 × SSC at about 65°C. Stringent hybridization conditions are hybridization conditions that are at least as stringent as the above representative conditions. Other stringent hybridization conditions are known in the art and may also be employed to identify nucleic acids of this particular embodiment of the invention.

Also provided are nucleic acids that encode a polypeptide having mono and/or diacylglycerol acyltransferase activity and having at least about 50%, at least about 60%, at least about 70&, at least about 75%, at least about 80%, at least about 90%, at least about 95%, or higher, nucleotide sequence identity to a nucleic acid having a nucleic acid sequence set forth in any one of SEQ ID NO:01, 03, 05, 07, 09, 11, 13, 15, or 18. Also provided are nucleic acids that encode a polypeptide having mono- and/or diacylglycerol acyltransferase activity and having at least about 50%, at least about 60%, at least about 70&, at least about 75%, at least about 80%, at least about 90%, at least about 95%, or higher, nucleotide sequence identity to the coding region of a nucleic acid having a sequence set forth in any one of SEQ ID NO:01, 03, 05, 07, 09, 11, 13, 15, or 18.

Nucleic acids encoding the DGAT2α proteins, DGAT2α polypeptides, and MGAT polypeptides of the subject invention may be cDNAs or genomic DNAs, i.e. portions of chromosomes that include both introns and exons, as well as promoter regions, etc., as well as fragments thereof. The term " DGAT2α - gene" shall be intended to mean the open reading frame encoding specific DGAT2α proteins and polypeptides, and DGAT2α introns, as well as adjacent 5′ and 3′ non-coding nucleotide sequences involved in the regulation of

expression, up to about 20 kb beyond the coding region, but possibly further in either direction. Similarly, the term "MGAT1 gene" refers to the open reading from encoding specific MGAT polypeptides, and MGAT1 introns, as well as adjacent 5' and 3' non-coding nucleotide sequences involved in the regulation of expression, up to about 20 kb beyond the coding region, but possibly further in either direction. The gene may be introduced into an appropriate vector for extrachromosomal maintenance or for integration into a host genome.

The term "cDNA" as used herein is intended to include all nucleic acids that share the arrangement of sequence elements found in native mature mRNA species, where sequence elements are exons and 3′ and 5′ non-coding regions. Normally mRNA species have contiguous exons, with the intervening introns, when present, being removed by nuclear RNA splicing, to create a continuous open reading frame encoding an MGAT1 or a DGAT2α protein.

Also provided are nucleic acids that encode the DGAT2α and MGAT1 proteins encoded by the above described nucleic acids, but differ in sequence from the above described nucleic acids due to the degeneracy of the genetic code. Also provided are nucleic acids that encode DGAT2α and MGAT1 proteins that include conservative amino acid changes when compared to, e.g., the amino acid sequences set forth in SEQ ID NO:02 or 06.

A genomic sequence of interest comprises the nucleic acid present between the initiation codon and the stop codon, as defined in the listed sequences, including all of the introns that are normally present in a native chromosome. It may further include the 3′ and 5′ untranslated regions found in the mature mRNA. It may further include specific transcriptional and translational regulatory sequences, such as promoters, enhancers, *etc.*, including about 1 kb, but possibly more, of flanking genomic DNA at either the 5′ or 3′ end of the transcribed region. The genomic DNA may be isolated as a fragment of 100 kbp or smaller; and substantially free of flanking chromosomal sequence. The genomic DNA flanking the coding region, either 3' or 5', or internal regulatory sequences as sometimes found in introns, contains sequences required for proper tissue and stage specific expression.

The nucleic acid compositions of the subject invention may encode all or a part of the subject DGAT2α or MGAT1 proteins and polypeptides, described in greater detail *infra*. Double or single stranded fragments may be obtained from the DNA sequence by chemically

synthesizing oligonucleotides in accordance with conventional methods, by restriction enzyme digestion, by PCR amplification, *etc.* For the most part, DNA fragments will be of at least 15 nt, usually at least 18 nt or 25 nt, and may be at least about 50 nt.

The MGAT1 and DGAT2α - nucleic acids or genes of the subject invention are isolated and obtained in substantial purity, generally as other than an intact chromosome. Usually, the DNA will be obtained substantially free of other nucleic acid sequences that do not include a DGAT2α sequence or fragment thereof, generally being at least about 50%, usually at least about 90% pure and are typically "recombinant", *i.e.* flanked by one or more nucleotides with which it is not normally associated on a naturally occurring chromosome.

In addition to the plurality of uses described in greater detail in following sections, the subject nucleic acid compositions find use in the preparation of all or a portion of the DGAT2α or MGAT1 polypeptides, as described below.

POLYPEPTIDE COMPOSITIONS

Also provided by the subject invention are polypeptides having mono- and/or diglyceride acyltransferase activity, i.e., capable of catalyzing the acylation of diacylglycerol, acylation of monoacylglycerol, or acylation of both mono- and diacylglycerol. Such enzymes are referred to herein as "mono- and diacylglcerol acyltransferases." Examples of such polypeptides are DGAT2α (also referred to as "DGAT2") and MGAT1 (also referred to as "DC2"). The term "polypeptide composition" as used herein refers to both full-length proteins as well as portions or fragments thereof. Also included in this term are variations of the naturally occurring proteins, where such variations are homologous or substantially similar to the naturally occurring protein, as described in greater detail below, be the naturally occurring protein the human protein, mouse protein, or protein from some other mammalian species which naturally expresses a subject acyltransferase. In the following description of the subject invention, the term "DGAT2α" is used to refer not only to the human form of the enzyme, but also to homologs thereof expressed in non-human mammalian species. Similarly, the term "mouse MGAT1" refers not only to the mouse form of the enzyme, but also to homologs thereof expressed in other mammalian species.

The subject mono- and diacylglcerol acyltransferases are, in their natural environment,

trans-membrane proteins. The subject proteins are characterized by the presence of at least one potential N-linked glycosylation site, at least one potential tyrosine phosphorylation site, and multiple hydrophobic domains, including 4 to 12, e.g., 6, hydrophobic domains capable of serving as trans-membrane regions. The proteins range in length from about 300 to 500, usually from about 325 to 475 and more usually from about 350 to 425 amino acid residues, and the projected molecular weight of the subject proteins based solely on the number of amino acid residues in the protein ranges from about 35 to 55, usually from about 37.5 to 47.5 and more usually from about 40 to 45 kDa, where the actual molecular weight may vary depending on the amount of glycolsylation of the protein and the apparent molecular weight may be considerably less because of SDS binding on gels.

The amino acid sequences of the subject proteins are characterized by having substantially no homology to the known DGAT enzymes. More specifically, the subject human DGAT2α and mouse MGAT1 enzymes have substantially no homology to the human DGAT enzyme described in Cases et al., "Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis," Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13018-13023 (1998). Likewise, the subject mouse DGAT2α enzymes have substantially no homology to the mouse DGAT enzyme described in Cases et al., "Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis," Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13018-13023 (1998). By substantially no homology is meant that the homology does not exceed about 20%, and usually will not exceed about 10% and more usually will not exceed about as determined using GCG (Genetics Computer Group, Wisconsin Package, Standard Settings, Gap Creation Penalty 3.0, Gap Extension Penalty 0.1).

Of particular interest in many embodiments are proteins that are non-naturally glycosylated. By non-naturally glycosylated is meant that the protein has a glycosylation pattern, if present, which is not the same as the glycosylation pattern found in the corresponding naturally occurring protein. For example, human DGAT2α of the subject invention and of this particular embodiment is characterized by having a glycosylation pattern, if it is glycosylated at all, that differs from that of naturally occurring human DGAT2α. Thus, the non-naturally glycosylated DGAT2α proteins of this embodiment

include non-glycosylated DGAT2α proteins, i.e. proteins having no covalently bound glycosyl groups.

The sequence of the full-length human DGAT2α protein is identified, infra, as SEQ ID NO:02. As such, DGAT2α proteins having an amino acid sequence that is substantially the same as or identical to the sequence of SEQ ID NO:2 are of interest. By substantially the same as is meant a protein having a region with a sequence that has at least about 75%, usually at least about 90% and more usually at least about 98 % sequence identity with the sequence of SED ID NO:02, as measured by GCG, supra. Of particular interest in other embodiments is the mouse DGAT2α protein, where the mouse DGAT2α protein of the subject invention has an amino acid sequence that is substantially the same as or identical to the sequence appearing as SEQ ID NO:04, infra.

In addition to the specific mammalian DGAT2α proteins described above, homologs or proteins (or fragments thereof) from other species, i.e. other animal or plant species, are also provided, where such homologs or proteins may be from a variety of different types of species, including animals, such as mammals, e.g., rodents, such as rats, mice; domestic animals, e.g. horse, cow, dog, cat; humans, and the like. By homolog is meant a protein having at least about 35 %, usually at least about 40% and more usually at least about 60 % amino acid sequence identity the specific DGAT2α proteins as identified in SEQ ID NOS: 02 to 04, where sequence identity is determined using GCG, *supra*. Specific homologs of interest include human DC 2, human DC3, human DC4, human DC5, mouse DC2 and mouse DC3, the sequences of which are provided in Fig. 8 (i.e., SEQ ID NOs. 06, 08, 10, 12, 14 and 16).

Mouse MGAT1 exhibits monoacylglycerol acyltransferase activity. The sequence of the full-length mouse MGAT1 protein is identified, infra, as SEQ ID NO:06 (identified as "mouse DC2" in Figure 8). As such, subject MGAT1 proteins having an amino acid sequence that is substantially the same as or identical to the sequence of SEQ ID NO:06 are of interest. By substantially the same as is meant a protein having a region with a sequence that has at least about 75%, usually at least about 90% and more usually at least about 98 % sequence identity with the sequence of SED ID NO:06, as measured by GCG, supra.

Mono- and diacylglcerol acyltransferases of the subject invention (e.g. human DGAT2α or a homolog thereof; non-human DGAT2α proteins, e.g. mouse DGAT2α; mouse

MGAT1 polypeptide or a homolog thereof) are present in a non-naturally occurring environment, *e.g.* are separated from their naturally occurring environment. In certain embodiments, the subject mono- and diacylglcerol acyltransferases are present in a composition that is enriched for such an enzyme, e.g., enriched for DGAT2α as compared to DGAT2α in its naturally occurring environment. As such, purified mono- and diacylglcerol acyltransferases are provided, where by purified is meant that subject enzyme is present in a composition that is substantially free of proteins other than the subject enzyme, where by substantially free is meant that less than 90 %, usually less than 60 % and more usually less than 50 % of the composition is made up of proteins other than the subject enzyme. For example, for compositions that are enriched for DGAT2α proteins, such compositions will exhibit a DGAT2α activity of at least about 100, usually at least about 200 and more usually at least about 1000 pmol triglycerides formed/mg protein/min, where such activity is determined by the assay described in the Experimental Section, infra.

In certain embodiments of interest, a subject enzyme is present in a composition that is substantially free of the constituents that are present in its naturally occurring environment. For example, a human DGAT2α protein comprising composition according to the subject invention in this embodiment will be substantially, if not completely, free of those other biological constituents, such as proteins, carbohydrates, lipids, etc., with which it is present in its natural environment. As such, protein compositions of these embodiments will necessarily differ from those that are prepared by purifying the protein from a naturally occurring source, where at least trace amounts of the protein's constituents will still be present in the composition prepared from the naturally occurring source.

The mono- and diacylglcerol acyltransferases of the subject invention may also be present as an isolate, by which is meant that the subject enzyme is substantially free of both proteins other than a subject enzyme and other naturally occurring biologic molecules, such as oligosaccharides, polynucleotides and fragments thereof, and the like, where substantially free in this instance means that less than 70 %, usually less than 60% and more usually less than 50 % (dry weight) of the composition containing the isolated subject enzyme is a naturally occurring biological molecule other than the subject enzyme. In certain embodiments, the subject enzyme is present in substantially pure form, where by substantially pure form is

meant at least 95%, usually at least 97% and more usually at least 99% pure.

In addition to the naturally occurring subject proteins, mono- and diacylglcerol acyltransferase polypeptides which vary from the naturally occurring DGAT2α and/or MGAT1 proteins are also provided. By "DGAT2α polypeptides" and "MGAT1 polypeptides" is meant proteins having an amino acid sequence encoded by an open reading frame (ORF) of a DGAT2α gene or an MGAT1 gene, respectively, as described *supra*, including the full length DGAT2α or MGAT1 protein and fragments thereof, particularly biologically active fragments and/or fragments corresponding to functional domains; and including fusions of the subject polypeptides to other proteins or parts thereof. Fragments of interest will typically be at least about 10 amino acids (aa) in length, usually at least about 50 aa in length, and may be as long as 300 aa in length or longer, but will usually not exceed about 1000 aa in length, where the fragment will have a stretch of amino acids that is identical to a subject protein of SEQ ID NO:2, SEQ ID NO:04, or SEQ ID NO:06 or a homolog thereof; of at least about 10 aa, and usually at least about 15 aa, and in many embodiments at least about 50 aa in length.

PREPARATION OF SUBJECT POLYPEPTIDES

The subject proteins and polypeptides may be obtained from naturally occurring sources, but are preferably synthetically produced. Where obtained from naturally occurring sources, the source chosen will generally depend on the species from which the subject protein is to be derived.

The subject polypeptide compositions may be synthetically derived by expressing a recombinant gene encoding the subject protein, such as the polynucleotide compositions described above, in a suitable host. For expression, an expression cassette may be employed. The expression vector will provide a transcriptional and translational initiation region, which may be inducible or constitutive, where the coding region is operably linked under the transcriptional control of the transcriptional initiation region, and a transcriptional and translational termination region. These control regions may be native to a DGAT2α gene, an MGAT1 gene, or may be derived from exogenous sources.

Expression vectors generally have convenient restriction sites located near the

promoter sequence to provide for the insertion of nucleic acid sequences encoding heterologous proteins. A selectable marker operative in the expression host may be present. Expression vectors may be used for the production of fusion proteins, where the exogenous fusion peptide provides additional functionality, i.e. increased protein synthesis, stability, reactivity with defined antisera, an enzyme marker, e.g. β-galactosidase, etc.

Expression cassettes may be prepared comprising a transcription initiation region, the gene or fragment thereof, and a transcriptional termination region. Of particular interest is the use of sequences that allow for the expression of functional epitopes or domains, usually at least about 8 amino acids in length, more usually at least about 15 amino acids in length, to about 25 amino acids, and up to the complete open reading-frame of the gene. After introduction of the DNA, the cells containing the construct may be selected by means of a selectable marker, the cells expanded and then used for expression.

Subject proteins and polypeptides may be expressed in prokaryotes or eukaryotes in accordance with conventional ways, depending upon the purpose for expression. For large scale production of the protein, a unicellular organism, such as *E. coli, B. subtilis, S. cerevisiae*, insect cells in combination with baculovirus vectors, or cells of a higher organism such as vertebrates, particularly mammals, *e.g.* COS 7 cells, may be used as the expression host cells. In some situations, it is desirable to express the subject coding sequence in eukaryotic cells, where the DGAT2α or MGAT1 protein will benefit from native folding and post-translational modifications. Small peptides can also be synthesized in the laboratory. Polypeptides that are subsets of the complete DGAT2α or MGAT1 sequence may be used to identify and investigate parts of the protein important for function.

Once the source of the protein is identified and/or prepared, e.g. a transfected host expressing the protein is prepared, the protein is then purified to produce the desired DGAT2α- or MGAT1- comprising composition. Any convenient protein purification procedures may be employed, where suitable protein purification methodologies are described in Guide to Protein Purification, (Deuthser ed.) (Academic Press, 1990). For example, a lysate may prepared from the original source, e.g. naturally occurring cells or tissues that express DGAT2α or the expression host expressing DGAT2α, and purified using HPLC, exclusion chromatography, gel electrophoresis, affinity chromatography, and the like.

Specific expression systems of interest include bacterial, yeast, insect cell and mammalian cell derived expression systems. Representative systems from each of these categories is are provided below:

Bacteria. Expression systems in bacteria include those described in Chang *et al.*, *Nature* (1978) *275*:615; Goeddel *et al.*, *Nature* (1979) *281*:544; Goeddel *et al.*, *Nucleic Acids Res.* (1980) *8*:4057; EP 0 036,776; U.S. Patent No. 4,551,433; DeBoer *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1983) *80*:21-25; and Siebenlist *et al.*, *Cell* (1980) *20*:269.

Yeast. Expression systems in yeast include those described in Hinnen *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1978) *75*:1929; Ito *et al.*, *J. Bacteriol.* (1983) *153*:163; Kurtz *et al.*, *Mol. Cell. Biol.* (1986) *6*:142; Kunze *et al.*, *J. Basic Microbiol.* (1985) *25*:141; Gleeson *et al.*, *J. Gen. Microbiol.* (1986) *132*:3459; Roggenkamp *et al.*, *Mol. Gen. Genet.* (1986) *202*:302; Das *et al.*, *J. Bacteriol.* (1984) *158*:1165; De Louvencourt *et al.*, *J. Bacteriol.* (1983) *154*:737; Van den Berg *et al.*, *Bio/Technology* (1990) *8*:135; Kunze *et al.*, *J. Basic Microbiol.* (1985) *25*:141; Cregg *et al.*, *Mol. Cell. Biol.* (1985) *5*:3376; U.S. Patent Nos. 4,837,148 and 4,929,555; Beach and Nurse, *Nature* (1981) *300*:706; Davidow *et al.*, *Curr. Genet.* (1985) *10*:380; Gaillardin *et al.*, *Curr. Genet.* (1985) *10*:49; Ballance *et al.*, *Biochem. Biophys. Res. Commun.* (1983) *112*:284-289; Tilburn *et al.*, *Gene* (1983) *26*:205-221; Yelton *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1984) *81*:1470-1474; Kelly and Hynes, *EMBO J.* (1985) *4*:475479; EP 0 244,234; and WO 91/00357.

Insect Cells. Expression of heterologous genes in insects is accomplished as described in U.S. Patent No. 4,745,051; Friesen *et al.*, "The Regulation of Baculovirus Gene Expression", in: *The Molecular Biology Of Baculoviruses* (1986) (W. Doerfler, ed.); EP 0 127,839; EP 0 155,476; and Vlak *et al.*, *J. Gen. Virol.* (1988) *69*:765-776; Miller *et al.*, *Ann. Rev. Microbiol.* (1988) *42*:177; Carbonell *et al.*, *Gene* (1988) *73*:409; Maeda *et al.*, *Nature* (1985) *315*:592-594; Lebacq-Verheyden *et al.*, *Mol. Cell. Biol.* (1988) *8*:3129; Smith *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1985) *82*:8844; Miyajima *et al.*, *Gene* (1987) *58*:273; and Martin *et al.*, *DNA* (1988) *7*:99. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts are described in Luckow *et al.*, *Bio/Technology* (1988) *6*:47-55, Miller *et al.*, *Generic Engineering* (1986) *8*:277-279, and Maeda *et al.*, *Nature* (1985) *315*:592-594.

Mammalian Cells. Mammalian expression is accomplished as described in Dijkema *et al.*, *EMBO J.* (1985) *4*:761, Gorman *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1982) *79*:6777, Boshart *et al.*, *Cell* (1985) *41*:521 and U.S. Patent No. 4,399,216. Other features of mammalian expression are facilitated as described in Ham and Wallace, *Meth. Enz.* (1979) *58*:44, Barnes and Sato, *Anal. Biochem.* (1980) *102*:255, U.S. Patent Nos. 4,767,704, 4,657,866, 4,927,762, 4,560,655, WO 90/103430, WO 87/00195, and U.S. RE 30,985.

When any of the above host cells, or other appropriate host cells or organisms, are used to replicate and/or express the polynucleotides or nucleic acids of the invention, the resulting replicated nucleic acid, RNA, expressed protein or polypeptide, is within the scope of the invention as a product of the host cell or organism. The product is recovered by any appropriate means known in the art.

Once the gene corresponding to a selected polynucleotide is identified, its expression can be regulated in the cell to which the gene is native. For example, an endogenous gene of a cell can be regulated by an exogenous regulatory sequence as disclosed in U.S. Patent No. 5,641,670.

METHODS AND COMPOSITIONS HAVING RESEARCH APPLICATION

Also provided by the subject invention are methods and compositions having research applications, such as in the study of the acylglycerol metabolism; in the identification of key components of the di- and triglyceride synthesis pathway; in the identification of di- and triglyceride synthesis modulatory agents, e.g. DGAT2α or MGAT1 inhibitors or enhancers, and the like.

The subject nucleic acid compositions find use in a variety of research applications. Research applications of interest include: the identification of DGAT2α and MGAT1 homologs; as a source of novel promoter elements; the identification of DGAT2α or MGAT1 expression regulatory factors; as probes and primers in hybridization applications, e.g. PCR; the identification of expression patterns in biological specimens; the preparation of cell or animal models for DGAT2α or MGAT1 function; the preparation of *in vitro* models for DGAT2α or MGAT1 function; etc.

Homologs of the specifically disclosed subject nucleic acids are identified by any of a number of methods. A fragment of the provided cDNA may be used as a hybridization probe against a cDNA library from the target organism of interest, where low stringency conditions are used. The probe may be a large fragment, or one or more short degenerate primers. Nucleic acids having sequence similarity are detected by hybridization under low stringency conditions, for example, at 50°C and 6×SSC (0.9 M sodium chloride/0.09 M sodium citrate) and remain bound when subjected to washing at 55°C in 1×SSC (0.15 M sodium chloride/0.015 M sodium citrate). Sequence identity may be determined by hybridization under stringent conditions, for example, at 50°C or higher and 0.1×SSC (15 mM sodium chloride/01.5 mM sodium citrate). Nucleic acids having a region of substantial identity to the provided nucleic acid sequences bind to the provided sequences under stringent hybridization conditions. By using probes, particularly labeled probes of DNA sequences, one can isolate homologous or related genes. One can also use sequence information derived from the polynucleotide compositions of the subject invention to prepare electronic "probes" for use in searching of computer based sequence date, e.g. BLAST searches EST databases.

The sequence of the 5' flanking region of the subject nucleic acid compositions may be utilized as a source for promoter elements, including enhancer-binding sites, that provide for developmental regulation in tissues where a subject acyltransferase, e.g., DGAT2α or MGAT1, is expressed. The tissue-specific expression is useful for determining the pattern of expression, and for providing promoters that mimic the native pattern of expression. Naturally occurring polymorphisms in the promoter region are useful for determining natural variations in expression, particularly those that may be associated with disease.

Alternatively, mutations may be introduced into the promoter region to determine the effect of altering expression in experimentally defined systems. Methods for the identification of specific DNA motifs involved in the binding of transcriptional factors are known in the art, *e.g.* sequence similarity to known binding motifs, gel retardation studies, *etc.* For examples, see Blackwell *et al.* (1995), *Mol. Med.* **1**:194-205; Mortlock *et al.* (1996), *Genome Res.* **6**:327-33; and Joulin and Richard-Foy (1995), *Eur. J. Biochem.* **232**:620-626.

The regulatory sequences may be used to identify *cis* acting sequences required for transcriptional or translational regulation of DGAT2 gene expression, especially in different

tissues or stages of development, and to identify *cis* acting sequences and *trans*-acting factors that regulate or mediate DGAT2 or MGAT1 gene expression. Such transcription or translational control regions may be operably linked to a DGAT2 or MGAT gene in order to promote expression of wild type or altered DGAT2 or MGAT1 or other proteins of interest in cultured cells, or in embryonic, fetal or adult tissues, and for gene therapy.

Small DNA fragments are useful as primers for PCR, hybridization screening probes, *etc.* Larger DNA fragments, *i.e.* greater than 100 nucleotides (nt) are useful for production of the encoded polypeptide, as described in the previous section. For use in amplification reactions, such as PCR, a pair of primers will be used. The exact composition of the primer sequences is not critical to the invention, but for most applications the primers will hybridize to the subject sequence under stringent conditions, as known in the art. It is preferable to choose a pair of primers that will generate an amplification product of at least about 50 nt, preferably at least about 100 nt. Algorithms for the selection of primer sequences are generally known, and are available in commercial software packages. Amplification primers hybridize to complementary strands of DNA, and will prime towards each other.

The DNA may also be used to identify expression of the gene in a biological specimen. The manner in which one probes cells for the presence of particular nucleotide sequences, as genomic DNA or RNA, is well established in the literature. Briefly, DNA or mRNA is isolated from a cell sample. The mRNA may be amplified by RT-PCR, using reverse transcriptase to form a complementary DNA strand, followed by polymerase chain reaction amplification using primers specific for the subject DNA sequences. Alternatively, the mRNA sample is separated by gel electrophoresis, transferred to a suitable support, *e.g.* nitrocellulose, nylon, *etc.*, and then probed with a fragment of the subject DNA as a probe. Other techniques, such as oligonucleotide ligation assays, *in situ* hybridizations, and hybridization to DNA probes arrayed on a solid chip may also find use. Detection of mRNA hybridizing to the subject sequence is indicative of DGAT2α or MGAT1 gene expression in the sample.

The sequence of a subject gene or nucleic acid, including flanking promoter regions and coding regions, may be mutated in various ways known in the art to generate targeted changes in promoter strength, sequence of the encoded protein, *etc.* The DNA sequence or

protein product of such a mutation will usually be substantially similar to the sequences provided herein, *i.e.* will differ by at least one nucleotide or amino acid, respectively, and may differ by at least two but not more than about ten nucleotides or amino acids. The sequence changes may be substitutions, insertions, deletions, or a combination thereof. Deletions may further include larger changes, such as deletions of a domain or exon. Other modifications of interest include epitope tagging, *e.g.* with the FLAG system, HA, *etc.* For studies of subcellular localization, fusion proteins with green fluorescent proteins (GFP) may be used.

Techniques for *in vitro* mutagenesis of cloned genes are known. Examples of protocols for site specific mutagenesis may be found in Gustin *et al. (1993)*, *Biotechniques* **14**:22; Barany (1985), *Gene* **37**:111-23; Colicelli *et al. (1985)*, *Mol. Gen. Genet.* **199**:537-9; and Prentki *et al.* (1984), *Gene* **29**:303-13. Methods for site specific mutagenesis can be found in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, CSH Press 1989, pp. 15.3-15.108; Weiner *et al.* (1993), *Gene* **126**:35-41; Sayers *et al.* (1992), *Biotechniques* **13**:592-6; Jones and Winistorfer (1992), *Biotechniques* **12**:528-30; Barton *et al.* (1990), *Nucleic Acids Res* **18**:7349-55; Marotti and Tomich (1989), *Gene Anal. Tech.* **6**:67-70; and Zhu (1989), *Anal Biochem* **177**:120-4. Such mutated genes may be used to study structure-function relationships of DGAT2α, or to alter properties of the protein that affect its function or regulation.

The subject nucleic acids can be used to generate transgenic hosts, e.g non-human animals, such as mice, cows, rats, pigs etc., or site specific gene modifications in cell lines. Examples of transgenic hosts include hosts in which the naturally expressed DGAT2α or MGAT1 gene has been disrupted, e.g. DGAT2α or MGAT1 knock-outs, as well as hosts in which DGAT2α or MGAT1 expression has been amplified, e.g. through introduction of additional DGAT2α or MGAT1 copies, through introduction of strong promoter upstream of the DGAT2α or MGAT1 gene, and the like. Using the nucleic acid compositions of the subject invention, standard protocols known to those of skill in the art may used to produce such transgenic hosts that have been genetically manipulated with respect to the subject gene, i.e. DGAT2α or MGAT1 transgenic hosts.

Transgenic animals may be made through homologous recombination, where the

normal DGAT2α or MGAT1 locus is altered, e.g. as in DGAT2α or MGAT1 knockouts. Alternatively, a nucleic acid construct is randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other animal viruses, YACs, and the like. DNA constructs for homologous recombination will comprise at least a portion of the DGAT2α or MGAT1 gene native to the species of the host animal, wherein the gene has the desired genetic modification(s), and includes regions of homology to the target locus. DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For various techniques for transfecting mammalian cells, see Keown *et al.* (1990), *Meth. Enzymol.* **185**:527-537.

For embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, *e.g.* mouse, rat, guinea pig, cow, *etc.* Such cells are grown on an appropriate fibroblast-feeder layer or grown in the presence of leukemia inhibiting factor (LIF). When ES or embryonic cells have been transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting offspring screened for the construct. By providing for a different phenotype of the blastocyst and the genetically modified cells, chimeric progeny can be readily detected.

The resultant chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogeneic or congenic grafts or transplants, or in *in vitro* culture. The

transgenic animals may be any non-human mammal, such as laboratory animals, domestic animals, etc.

Transgenic plants may be produced in a similar manner. Methods of preparing transgenic plant cells and plants are described in U.S. Pat. Nos. 5,767,367; 5,750,870; 5,739,409; 5,689,049; 5,689,045; 5,674,731; 5,656,466; 5,633,155; 5,629,470 ; 5,595,896; 5,576,198; 5,538,879; 5,484,956; the disclosures of which are herein incorporated by reference. Methods of producing transgenic plants are also reviewed in Plant Biochemistry and Molecular Biology (eds Lea & Leegood, John Wiley & Sons)(1993) pp 275-295. In brief, a suitable plant cell or tissue is harvested, depending on the nature of the plant species. As such, in certain instances, protoplasts will be isolated, where such protoplasts may be isolated from a variety of different plant tissues, e.g. leaf, hypoctyl, root, etc. For protoplast isolation, the harvested cells are incubated in the presence of cellulases in order to remove the cell wall, where the exact incubation conditions vary depending on the type of plant and/or tissue from which the cell is derived. The resultant protoplasts are then separated from the resultant cellular debris by sieving and centrifugation. Instead of using protoplasts, embryogenic explants comprising somatic cells may be used for preparation of the transgenic host. Following cell or tissue harvesting, exogenous DNA of interest is introduced into the plant cells, where a variety of different techniques are available for such introduction. With isolated protoplasts, the opportunity arise for introduction via DNA-mediated gene transfer protocols, including: incubation of the protoplasts with naked DNA, e.g. plasmids, comprising the exogenous coding sequence of interest in the presence of polyvalent cations, e.g. PEG or PLO; and electroporation of the protoplasts in the presence of naked DNA comprising the exogenous sequence of interest. Protoplasts that have successfully taken up the exogenous DNA are then selected, grown into a callus, and ultimately into a transgenic plant through contact with the appropriate amounts and ratios of stimulatory factors, e.g. auxins and cytokinins. With embryogenic explants, a convenient method of introducing the exogenous DNA in the target somatic cells is through the use of particle acceleration or "gene-gun" protocols. The resultant explants are then allowed to grow into chimera plants, cross-bred and transgenic progeny are obtained. Instead of the naked DNA approaches described above, another convenient method of producing transgenic plants is *Agrobacterium* mediated

transformation. With *Agrobacterium* mediated transformation, co-integrative or binary vectors comprising the exogenous DNA are prepared and then introduced into an appropriate *Agrobacterium* strain, e.g. *A. tumefaciens*. The resultant bacteria are then incubated with prepared protoplasts or tissue explants, e.g. leaf disks, and a callus is produced. The callus is then grown under selective conditions, selected and subjected to growth media to induce root and shoot growth to ultimately produce a transgenic plant.

The modified cells, animals or plants are useful in the study of function and regulation or a subject gene or nucleic acid. For example, a series of small deletions and/or substitutions may be made in the host's native DGAT2α or MGAT1 gene to determine the role of different exons in various physiological processes. Specific constructs of interest include anti-sense nucleic acids, which will block DGAT2α or MGAT1 expression, expression of dominant negative DGAT2α or MGAT1 mutations, and over-expression of DGAT2α or MGAT1 genes. Where a subject nucleic acid sequence is introduced, the introduced sequence may be either a complete or partial sequence of a subject gene native to the host, or may be a complete or partial subject nucleic acid sequence that is exogenous to the host animal, *e.g.*, a human DGAT2α or a human MGAT1 sequence. A detectable marker, such as *lac Z* (encoding β-galactosidase) may be introduced into the DGAT2α or MGAT1 locus, where upregulation of DGAT2α or MGAT1 gene expression will result in an easily detected change in phenotype. One may also provide for expression of the subject gene or variants thereof in cells or tissues where it is not normally expressed, at levels not normally present in such cells or tissues, or at abnormal times of development. The transgenic hosts, e.g. animals, plants, etc., may be used in functional studies, drug screening, *etc.*, *e.g.* to determine the effect of a candidate drug on DGAT2α or MGAT1 activity, to identify drugs that reduce serum triglyceride levels, etc.

The subject polypeptide compositions can be used to produce *in vitro* models of diglyceride and/or triglyceride synthesis, where such models will consist of the subject proteins and other components of di- and/or triglyceride synthesis, e.g. substrates, such as monoacylglycerol, diacylglycerol or metabolic precursors thereof, fatty acyl CoAs and the like, other components of the diacylglycerol and/or triacylglycerol synthetase complex, e.g. acyl CoA ligase, acyl CoA acyltransferase, monoacyl glycerol acyltransferase, etc.

Also provided by the subject invention are screening assays designed to find modulatory agents of activity of a subject mono- or diacylglycerol acyltransferase, e.g. inhibitors or enhancers of DGAT2α or MGAT1 activity, as well as the agents identified thereby, where such agents may find use in a variety of applications, including as therapeutic agents, as agricultural chemicals, etc. The screening methods will typically be assays which provide for qualitative/quantitative measurements of DGAT2α or MGAT1 activity in the presence of a particular candidate therapeutic agent. For example, the assay could be an assay which measures the acylation activity of DGAT2α or MGAT1 in the presence and absence of a candidate inhibitor agent. The screening method may be an *in vitro* or *in vivo* format, where both formats are readily developed by those of skill in the art.

Thus, in some embodiments, the invention provides an *in vitro* method of identifying an agent that modulates the acyltransferase activity of MGAT1. The method generally involves contacting MGAT1 with a candidate agent (also referred to as a "test agent") in the presence of an acyl donor and an acyl acceptor. The effect, if any, of a test agent on the amount of acylated acceptor that is produced is measured relative to a control sample, which control sample includes the MGAT1 polypeptide, the acyl donor, and the acyl acceptor, and no test agent. Typically, the reaction mixture includes magnesium ions (e.g., $MgCl_2$); and a buffer. Exemplary reaction conditions are provided in Example 5. Suitable acyl donors are fatty acyl CoA compounds and include, but are not limited to, oleoyl CoA. Typically, the acyl group of the acyl donor is labeled with a detectable label, such that when the acyl group is transferred to the acyl acceptor, the detectable label is also transferred, thereby allowing detection of the acylated acceptor molecule. Suitable acyl acceptors are monoacylglyerols and diacylglycerols.

The DGAT2α or MGAT1 polypeptide in the screening assay may be purified, but need not be. For example, membrane fractions containing DGAT2α or MGAT1 polypeptides can be used.

Depending on the particular method, one or more of, usually one of, the components of the screening assay may be labeled, where by labeled is meant that the components comprise a detectable moiety, *e.g.* a fluorescent or radioactive tag, or a member of a signal producing system, *e.g.* biotin for binding to an enzyme-streptavidin conjugate in which the

enzyme is capable of converting a substrate to a chromogenic product. Where in vitro assays are employed, the various components of the in vitro assay, e.g. the substrate, the donor, the DGAT2α or MGAT1 protein and the candidate inhibitor, etc. are combined in a assay mixture under conditions sufficient for DGAT2α or MGAT1 activity to occur, as described in the experimental section, *infra*.

A variety of other reagents may be included in the screening assay and reaction mixture. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc that are used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Reagents that improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, *etc.* may be used.

A variety of different candidate agents may be screened by the above methods. Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification,

amidification, *etc.* to produce structural analogs.

Using the above screening methods, a variety of different therapeutic agents may be identified. Such agents may target the enzyme itself, or an expression regulator factor thereof. Such agents may inhibitors or promoters of DGAT2α or MGAT1 activity, where inhibitors are those agents that result in at least a reduction of DGAT2α or MGAT1 activity as compared to a control and enhancers result in at least an increase in DGAT2α or MGAT1 activity as compared to a control. Such agents may be find use in a variety of therapeutic applications, as described in greater detail below.

METHODS AND COMPOSITIONS HAVING MEDICAL APPLICATIONS

The methods and compositions of the subject invention also have broad ranging applications in a variety of medical applications, including diagnostic screening, therapeutic treatments of pathological conditions, in the regulation of DGAT2α or MGAT1 activity in desirable ways, and the like.

The subject invention provides methods of screening individuals for a predisposition to a disease state or the presence of disease state, where such screening may focus on the presence of one or more markers, such as a mutated DGAT2α or MGAT1 gene or expression regulatory element thereof, observed levels of DGAT2α or MGAT1; the expression level of the DGAT2α or MGAT1 gene in a biological sample of interest; and the like.

Samples, as used herein, include biological fluids such as blood, cerebrospinal fluid, tears, saliva, lymph, semen, dialysis fluid and the like; organ or tissue culture derived fluids; and fluids extracted from physiological tissues. Also included in the term are derivatives and fractions of such fluids. The cells may be dissociated, in the case of solid tissues, or tissue sections may be analyzed. Alternatively a lysate of the cells may be prepared.

A number of methods are available for determining the expression level of a gene or protein in a particular sample. Diagnosis may be performed by a number of methods to determine the absence or presence or altered amounts of normal or abnormal DGAT2α or MGAT1 in a patient sample. For example, detection may utilize staining of cells or

histological sections with labeled antibodies, performed in accordance with conventional methods. Cells are permeabilized to stain cytoplasmic molecules. The antibodies of interest are added to the cell sample, and incubated for a period of time sufficient to allow binding to the epitope, usually at least about 10 minutes. The antibody may be labeled with radioisotopes, enzymes, fluorescers, chemiluminescers, or other labels for direct detection. Alternatively, a second stage antibody or reagent is used to amplify the signal. Such reagents are well known in the art. For example, the primary antibody may be conjugated to biotin, with horseradish peroxidase-conjugated avidin added as a second stage reagent. Alternatively, the secondary antibody conjugated to a flourescent compound, *e.g.* fluorescein, rhodamine, Texas red, *etc.* Final detection uses a substrate that undergoes a color change in the presence of the peroxidase. The absence or presence of antibody binding may be determined by various methods, including flow cytometry of dissociated cells, microscopy, radiography, scintillation counting, *etc.*

Alternatively, one may focus on the expression of DGAT2α- or MGAT1- encoding nucleic acids. Biochemical studies may be performed to determine whether a sequence polymorphism in a DGAT2α or MGAT1 coding region or control regions is associated with disease. Disease associated polymorphisms may include deletion or truncation of the gene, mutations that alter expression level, that affect the activity of the protein, *etc.*

Changes in the promoter or enhancer sequence that may affect expression levels of DGAT2α or MGAT1 can be compared to expression levels of the normal allele by various methods known in the art. Methods for determining promoter or enhancer strength include quantitation of the expressed natural protein; insertion of the variant control element into a vector with a reporter gene such as β–galactosidase, luciferase, chloramphenicol acetyltransferase, *etc.* that provides for convenient quantitation; and the like.

A number of methods are available for analyzing nucleic acids for the presence of a specific sequence, *e.g.* a disease associated polymorphism. Where large amounts of DNA are available, genomic DNA is used directly. Alternatively, the region of interest is cloned into a suitable vector and grown in sufficient quantity for analysis. Cells that express DGAT2α or MGAT1 may be used as a source of mRNA, which may be assayed directly or reverse transcribed into cDNA for analysis. The nucleic acid may be amplified by conventional

techniques, such as the polymerase chain reaction (PCR), to provide sufficient amounts for analysis. The use of the polymerase chain reaction is described in Saiki, *et al.* (1985), *Science* **239**:487, and a review of techniques may be found in Sambrook, *et al.* Molecular Cloning: A Laboratory Manual, CSH Press 1989, pp.14.2-14.33. Alternatively, various methods are known in the art that utilize oligonucleotide ligation as a means of detecting polymorphisms, for examples see Riley *et al.* (1990), *Nucl. Acids Res.* **18**:2887-2890; and Delahunty *et al.* (1996), *Am. J. Hum. Genet.* **58**:1239-1246.

A detectable label may be included in an amplification reaction. Suitable labels include fluorochromes, *e.g.* fluorescein isothiocyanate (FITC), rhodamine, Texas Red, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein (JOE), 6-carboxy-X-rhodamine (ROX), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM) or N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA), radioactive labels, *e.g.* ^{32}P, ^{35}S, ^{3}H; *etc.* The label may be a two stage system, where the amplified DNA is conjugated to biotin, haptens, *etc.* having a high affinity binding partner, *e.g.* avidin, specific antibodies, *etc.*, where the binding partner is conjugated to a detectable label. The label may be conjugated to one or both of the primers. Alternatively, the pool of nucleotides used in the amplification is labeled, so as to incorporate the label into the amplification product.

The sample nucleic acid, *e.g.* amplified or cloned fragment, is analyzed by one of a number of methods known in the art. The nucleic acid may be sequenced by dideoxy or other methods, and the sequence of bases compared to a wild-type DGAT2α or MGAT1 sequence. Hybridization with the variant sequence may also be used to determine its presence, by Southern blots, dot blots, *etc.* The hybridization pattern of a control and variant sequence to an array of oligonucleotide probes immobilized on a solid support, as described in US 5,445,934, or in WO 95/35505, may also be used as a means of detecting the presence of variant sequences. Single strand conformational polymorphism (SSCP) analysis, denaturing gradient gel electrophoresis (DGGE), and heteroduplex analysis in gel matrices are used to detect conformational changes created by DNA sequence variation as alterations in electrophoretic mobility. Alternatively, where a polymorphism creates or destroys a recognition site for a restriction endonuclease, the sample is digested with that endonuclease,

and the products size fractionated to determine whether the fragment was digested. Fractionation is performed by gel or capillary electrophoresis, particularly acrylamide or agarose gels.

Screening for mutations in DGAT2α or MGAT1 may be based on the functional or antigenic characteristics of the protein. Protein truncation assays are useful in detecting deletions that may affect the biological activity of the protein. Various immunoassays designed to detect polymorphisms in DGAT2α or MGAT1 proteins may be used in screening. Where many diverse genetic mutations lead to a particular disease phenotype, functional protein assays have proven to be effective screening tools. The activity of the encoded DGAT2α or MGAT1 protein may be determined by comparison with the wild-type protein.

Diagnostic methods of the subject invention in which the level of DGAT2α or MGAT1 expression is of interest will typically involve comparison of the DGAT2α or MGAT1 nucleic acid abundance of a sample of interest with that of a control value to determine any relative differences, where the difference may be measured qualitatively and/or quantitatively, which differences are then related to the presence or absence of an abnormal DGAT2α or MGAT1 expression pattern. A variety of different methods for determining the nucleic acid abundance in a sample are known to those of skill in the art, where particular methods of interest include those described in: Pietu et al., Genome Res. (June 1996) 6: 492-503; Zhao et al., Gene (April 24, 1995) 156: 207-213; Soares , Curr. Opin. Biotechnol. (October 1997) 8: 542-546; Raval, J. Pharmacol Toxicol Methods (November 1994) 32: 125-127; Chalifour et al., Anal. Biochem (February 1, 1994) 216: 299-304; Stolz & Tuan, Mol. Biotechnol. (December 19960 6: 225-230; Hong et al., Bioscience Reports (1982) 2: 907; and McGraw, Anal. Biochem. (1984) 143: 298. Also of interest are the methods disclosed in WO 97/27317, the disclosure of which is herein incorporated by reference.

The subject diagnostic or screening methods may be used to identify the presence of, or predisposition to, disease conditions associated with acylglycerol metabolism, particularly those associated with DGAT2α or MGAT1 activity. Such disease conditions include: hyperlipidemia (including excess serum triglyceride levels), cardiovascular disease, obesity, diabetes, cancer, neurological disorders, immunological disorders, and the like.

Also provided are methods of regulating, including enhancing and inhibiting,

DGAT2α or MGAT1 activity in a host. A variety of situations arise where modulation of DGAT2α or MGAT1 activity in a host is desired, where such conditions include disease conditions associated with DGAT2α or MGAT1 activity and non-disease conditions in which a modulation of DGAT2α or MGAT1 activity is desired for a variety of different reasons, e.g. cosmetic weight control.

For the modulation of DGAT2α or MGAT1 activity in a host, an effective amount of active agent that modulates the activity, e.g. reduces the activity, of DGAT2α or MGAT1 *in vivo*, is administered to the host. The active agent may be a variety of different compounds, including: the polynucleotide compositions of the subject invention, the polypeptide compositions of the subject invention, a naturally occurring or synthetic small molecule compound, an antibody, fragment or derivative thereof, an antisense composition, and the like.

The nucleic acid compositions of the subject invention find use as therapeutic agents in situations where one wishes to enhance DGAT2α or MGAT1 activity in a host, e.g. in a mammalian host in which DGAT2α or MGAT1 activty is low resulting in a disease condition, etc. The DGAT2α or MGAT1 genes, gene fragments, or the encoded DGAT2α or MGAT1 protein or protein fragments are useful in gene therapy to treat disorders associated with DGAT2α or MGAT1 defects. Expression vectors may be used to introduce the DGAT2α gene or encoding nucleic acid into a cell. Such vectors generally have convenient restriction sites located near the promoter sequence to provide for the insertion of nucleic acid sequences. Transcription cassettes may be prepared comprising a transcription initiation region, the target gene or fragment thereof, and a transcriptional termination region. The transcription cassettes may be introduced into a variety of vectors, *e.g.* plasmid; retrovirus, *e.g.* lentivirus; adenovirus; and the like, where the vectors are able to transiently or stably be maintained in the cells, usually for a period of at least about one day, more usually for a period of at least about several days to several weeks.

Naturally occurring or synthetic small molecule compounds of interest include numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500

daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Of particular interest are those agents identified by the screening assays of the subject invention, as described above.

Also of interest as active agents are antibodies that modulate, e.g. reduce, if not inhibit, DGAT2α or MGAT1 activity in the host. Suitable antibodies are obtained by immunizing a host animal with peptides comprising all or a portion of a DGAT2 or MGAT1 protein, such as the DGAT2α or MGAT1 polypeptide compositions of the subject invention. Suitable host animals include mouse, rat sheep, goat, hamster, rabbit, *etc.* The origin of the protein immunogen may be mouse, human, rat, monkey *etc.* The host animal will generally be a different species than the immunogen, *e.g.* human DGAT2 used to immunize mice, *etc.*

The immunogen may comprise the complete protein, or fragments and derivatives thereof. Preferred immunogens comprise all or a part of DGAT2α or MGAT1, where these residues contain the post-translation modifications, such as glycosylation, found on the native DGAT2α or MGAT1. Immunogens comprising the extracellular domain are produced in a variety of ways known in the art, *e.g.* expression of cloned genes using conventional recombinant methods, isolation from HEC, *etc.*

For preparation of polyclonal antibodies, the first step is immunization of the host animal with DGAT2α of MGAT1, where the DGAT2α or MGAT1 protein will preferably be in substantially pure form, comprising less than about 1% contaminant. The immunogen may comprise complete DGAT2α or MGAT1, fragments or derivatives thereof. To increase the immune response of the host animal, the DGAT2 or MGAT1 may be combined with an adjuvant, where suitable adjuvants include alum, dextran, sulfate, large polymeric anions, oil & water emulsions, e.g. Freund's adjuvant, Freund's complete adjuvant, and the like. The DGAT2α or MGAT1 may also be conjugated to synthetic carrier proteins or synthetic

antigens. A variety of hosts may be immunized to produce the polyclonal antibodies. Such hosts include rabbits, guinea pigs, rodents, e.g. mice, rats, sheep, goats, and the like. The DGAT2α or MGAT1 is administered to the host, usually intradermally, with an initial dosage followed by one or more, usually at least two, additional booster dosages. Following immunization, the blood from the host will be collected, followed by separation of the serum from the blood cells. The Ig present in the resultant antiserum may be further fractionated using known methods, such as ammonium salt fractionation, DEAE chromatography, and the like.

Monoclonal antibodies are produced by conventional techniques. Generally, the spleen and/or lymph nodes of an immunized host animal provide a source of plasma cells. The plasma cells are immortalized by fusion with myeloma cells to produce hybridoma cells. Culture supernatant from individual hybridomas is screened using standard techniques to identify those producing antibodies with the desired specificity. Suitable animals for production of monoclonal antibodies to the human protein include mouse, rat, hamster, *etc.* To raise antibodies against the mouse protein, the animal will generally be a hamster, guinea pig, rabbit, etc. The antibody may be purified from the hybridoma cell supernatants or ascites fluid by conventional techniques, *e.g.* affinity chromatography using DGAT2α or MGAT1 bound to an insoluble support, protein A sepharose, *etc.*

The antibody may be produced as a single chain, instead of the normal multimeric structure. Single chain antibodies are described in Jost *et al.* (1994) J.B.C. **269**:26267-73, and others. DNA sequences encoding the variable region of the heavy chain and the variable region of the light chain are ligated to a spacer encoding at least about 4 amino acids of small neutral amino acids, including glycine and/or serine. The protein encoded by this fusion allows assembly of a functional variable region that retains the specificity and affinity of the original antibody.

For *in vivo* use, particularly for injection into humans, it is desirable to decrease the antigenicity of the antibody. An immune response of a recipient against the blocking agent will potentially decrease the period of time that the therapy is effective. Methods of humanizing antibodies are known in the art. The humanized antibody may be the product of an animal having transgenic human immunoglobulin constant region genes (see for example

International Patent Applications WO 90/10077 and WO 90/04036). Alternatively, the antibody of interest may be engineered by recombinant DNA techniques to substitute the CH1, CH2, CH3, hinge domains, and/or the framework domain with the corresponding human sequence (see WO 92/02190).

The use of Ig cDNA for construction of chimeric immunoglobulin genes is known in the art (Liu *et al.* (1987) P.N.A.S. **84**:3439 and (1987) J. Immunol. **139**:3521). mRNA is isolated from a hybridoma or other cell producing the antibody and used to produce cDNA. The cDNA of interest may be amplified by the polymerase chain reaction using specific primers (U.S. Patent nos. 4,683,195 and 4,683,202). Alternatively, a library is made and screened to isolate the sequence of interest. The DNA sequence encoding the variable region of the antibody is then fused to human constant region sequences. The sequences of human constant regions genes may be found in Kabat *et al.* (1991) Sequences of Proteins of Immunological Interest, N.I.H. publication no. 91-3242. Human C region genes are readily available from known clones. The choice of isotype will be guided by the desired effector functions, such as complement fixation, or activity in antibody-dependent cellular cytotoxicity. Preferred isotypes are IgG1, IgG3 and IgG4. Either of the human light chain constant regions, kappa or lambda, may be used. The chimeric, humanized antibody is then expressed by conventional methods.

Antibody fragments, such as Fv, $F(ab')_2$ and Fab may be prepared by cleavage of the intact protein, *e.g.* by protease or chemical cleavage. Alternatively, a truncated gene is designed. For example, a chimeric gene encoding a portion of the $F(ab')_2$ fragment would include DNA sequences encoding the CH1 domain and hinge region of the H chain, followed by a translational stop codon to yield the truncated molecule.

Consensus sequences of H and L J regions may be used to design oligonucleotides for use as primers to introduce useful restriction sites into the J region for subsequent linkage of V region segments to human C region segments. C region cDNA can be modified by site directed mutagenesis to place a restriction site at the analogous position in the human sequence.

Expression vectors include plasmids, retroviruses, YACs, EBV derived episomes, and

the like. A convenient vector is one that encodes a functionally complete human CH or CL immunoglobulin sequence, with appropriate restriction sites engineered so that any VH or VL sequence can be easily inserted and expressed. In such vectors, splicing usually occurs between the splice donor site in the inserted J region and the splice acceptor site preceding the human C region, and also at the splice regions that occur within the human CH exons. Polyadenylation and transcription termination occur at native chromosomal sites downstream of the coding regions. The resulting chimeric antibody may be joined to any strong promoter, including retroviral LTRs, *e.g.* SV-40 early promoter, (Okayama *et al.* (1983) Mol. Cell. Bio. **3**:280), Rous sarcoma virus LTR (Gorman *et al.* (1982) P.N.A.S. **79**:6777), and moloney murine leukemia virus LTR (Grosschedl *et al.* (1985) Cell **41**:885); native Ig promoters, *etc.*

In yet other embodiments of the invention, the active agent is an agent that modulates, and generally decreases or down regulates, the expression of DGAT2α- or MGAT1-encoding nucleic acids in the host. Antisense molecules can be used to down-regulate expression of these target nucleic acids in cells. The anti-sense reagent may be antisense oligonucleotides (ODN), particularly synthetic ODN having chemical modifications from native nucleic acids, or nucleic acid constructs that express such anti-sense molecules as RNA. The antisense sequence is complementary to the mRNA of the targeted gene, and inhibits expression of the targeted gene products. Antisense molecules inhibit gene expression through various mechanisms, *e.g.* by reducing the amount of mRNA available for translation, through activation of RNAse H, or steric hindrance. One or a combination of antisense molecules may be administered, where a combination may comprise multiple different sequences.

Antisense molecules may be produced by expression of all or a part of the target gene sequence in an appropriate vector, where the transcriptional initiation is oriented such that an antisense strand is produced as an RNA molecule. Alternatively, the antisense molecule is a synthetic oligonucleotide. Antisense oligonucleotides will generally be at least about 7, usually at least about 12, more usually at least about 20 nucleotides in length, and not more than about 500, usually not more than about 50, more usually not more than about 35 nucleotides in length, where the length is governed by efficiency of inhibition, specificity, including absence of cross-reactivity, and the like. It has been found that short oligonucleotides, of from 7 to 8 bases in length, can be strong and selective inhibitors of gene

expression (see Wagner *et al.* (1996), *Nature Biotechnol.* **14**:840-844).

A specific region or regions of the endogenous sense strand mRNA sequence is chosen to be complemented by the antisense sequence. Selection of a specific sequence for the oligonucleotide may use an empirical method, where several candidate sequences are assayed for inhibition of expression of the target gene in an *in vitro* or animal model. A combination of sequences may also be used, where several regions of the mRNA sequence are selected for antisense complementation.

Antisense oligonucleotides may be chemically synthesized by methods known in the art (see Wagner *et al.* (1993), *supra,* and Milligan *et al., supra.*) Preferred oligonucleotides are chemically modified from the native phosphodiester structure, in order to increase their intracellular stability and binding affinity. A number of such modifications have been described in the literature, which alter the chemistry of the backbone, sugars or heterocyclic bases.

Among useful changes in the backbone chemistry are phosphorothioates; phosphorodithioates, where both of the non-bridging oxygens are substituted with sulfur; phosphoroamidites; alkyl phosphotriesters and boranophosphates. Achiral phosphate derivatives include 3'-O'-5'-S-phosphorothioate, 3'-S-5'-O-phosphorothioate, 3'-CH2-5'-O-phosphonate and 3'-NH-5'-O-phosphoroamidate. Peptide nucleic acids replace the entire ribose phosphodiester backbone with a peptide linkage. Sugar modifications are also used to enhance stability and affinity. The α-anomer of deoxyribose may be used, where the base is inverted with respect to the natural β-anomer. The 2'-OH of the ribose sugar may be altered to form 2'-O-methyl or 2'-O-allyl sugars, which provides resistance to degradation without comprising affinity. Modification of the heterocyclic bases must maintain proper base pairing. Some useful substitutions include deoxyuridine for deoxythymidine; 5-methyl-2'-deoxycytidine and 5-bromo-2'-deoxycytidine for deoxycytidine. 5- propynyl-2'-deoxyuridine and 5-propynyl-2'-deoxycytidine have been shown to increase affinity and biological activity when substituted for deoxythymidine and deoxycytidine, respectively.

As an alternative to anti-sense inhibitors, catalytic nucleic acid compounds, *e.g.* ribozymes, anti-sense conjugates, *etc.* may be used to inhibit gene expression. Ribozymes may be synthesized *in vitro* and administered to the patient, or may be encoded on an

expression vector, from which the ribozyme is synthesized in the targeted cell (for example, see International patent application WO 9523225, and Beigelman *et al.* (1995), *Nucl. Acids Res.* **23**:4434-42). Examples of oligonucleotides with catalytic activity are described in WO 9506764. Conjugates of anti-sense ODN with a metal complex, *e.g.* terpyridylCu(II), capable of mediating mRNA hydrolysis are described in Bashkin *et al.* (1995), *Appl. Biochem. Biotechnol.* **54**:43-56.

As mentioned above, an effective amount of the active agent is administered to the host, where "effective amount" means a dosage sufficient to produce a desired result, where the desired result in the desired modulation, e.g. enhancement, reduction, of DGAT2α activity.

In the subject methods, the active agent(s) may be administered to the host using any convenient means capable of resulting in the desired effect. Thus, the agent can be incorporated into a variety of formulations for therapeutic administration. More particularly, the agents of the present invention can be formulated into pharmaceutical compositions by combination with appropriate, pharmaceutically acceptable carriers or diluents, and may be formulated into preparations in solid, semi-solid, liquid or gaseous forms, such as tablets, capsules, powders, granules, ointments, solutions, suppositories, injections, inhalants and aerosols.

As such, administration of the agents can be achieved in various ways, including oral, buccal, rectal, parenteral, intraperitoneal, intradermal, transdermal, intracheal,etc., administration.

In pharmaceutical dosage forms, the agents may be administered in the form of their pharmaceutically acceptable salts, or they may also be used alone or in appropriate association, as well as in combination, with other pharmaceutically active compounds. The following methods and excipients are merely exemplary and are in no way limiting.

For oral preparations, the agents can be used alone or in combination with appropriate additives to make tablets, powders, granules or capsules, for example, with conventional additives, such as lactose, mannitol, corn starch or potato starch; with binders, such as crystalline cellulose, cellulose derivatives, acacia, corn starch or gelatins; with disintegrators, such as corn starch, potato starch or sodium carboxymethylcellulose; with lubricants, such as

talc or magnesium stearate; and if desired, with diluents, buffering agents, moistening agents, preservatives and flavoring agents.

The agents can be formulated into preparations for injection by dissolving, suspending or emulsifying them in an aqueous or nonaqueous solvent, such as vegetable or other similar oils, synthetic aliphatic acid glycerides, esters of higher aliphatic acids or propylene glycol; and if desired, with conventional additives such as solubilizers, isotonic agents, suspending agents, emulsifying agents, stabilizers and preservatives.

The agents can be utilized in aerosol formulations to be administered via inhalation. The compounds of the present invention can be formulated into pressurized acceptable propellants such as dichlorodifluoromethane, propane, nitrogen and the like.

Furthermore, the agents can be made into suppositories by mixing with a variety of bases such as emulsifying bases or water-soluble bases. The compounds of the present invention can be administered rectally via a suppository. The suppository can include vehicles such as cocoa butter, carbowaxes and polyethylene glycols, which melt at body temperature, yet are solidified at room temperature.

Unit dosage forms for oral or rectal administration such as syrups, elixirs, and suspensions may be provided wherein each dosage unit, for example, teaspoonful, tablespoonful, tablet or suppository, contains a predetermined amount of the composition containing one or more inhibitors. Similarly, unit dosage forms for injection or intravenous administration may comprise the inhibitor(s) in a composition as a solution in sterile water, normal saline or another pharmaceutically acceptable carrier.

The term "unit dosage form," as used herein, refers to physically discrete units suitable as unitary dosages for human and animal subjects, each unit containing a predetermined quantity of compounds of the present invention calculated in an amount sufficient to produce the desired effect in association with a pharmaceutically acceptable diluent, carrier or vehicle. The specifications for the novel unit dosage forms of the present invention depend on the particular compound employed and the effect to be achieved, and the pharmacodynamics associated with each compound in the host.

The pharmaceutically acceptable excipients, such as vehicles, adjuvants, carriers or diluents, are readily available to the public. Moreover, pharmaceutically acceptable auxiliary

substances, such as pH adjusting and buffering agents, tonicity adjusting agents, stabilizers, wetting agents and the like, are readily available to the public.

Where the agent is a polypeptide, polynucleotide, analog or mimetic thereof, e.g. antisense composition, it may be introduced into tissues or host cells by any number of routes, including viral infection, microinjection, or fusion of vesicles. Jet injection may also be used for intramuscular administration, as described by Furth *et al.* (1992), *Anal Biochem* **205**:365-368. The DNA may be coated onto gold microparticles, and delivered intradermally by a particle bombardment device, or "gene gun" as described in the literature (see, for example, Tang *et al.* (1992), *Nature* **356**:152-154), where gold microprojectiles are coated with the *DGAT* DNA, then bombarded into skin cells.

Those of skill will readily appreciate that dose levels can vary as a function of the specific compound, the severity of the symptoms and the susceptibility of the subject to side effects. Preferred dosages for a given compound are readily determinable by those of skill in the art by a variety of means.

The subject methods find use in the treatment of a variety of different disease conditions involving acylglycerol metabolism, and particularly DGAT2α activity, including both insufficient or hypo- DGAT2α activity and hyper- DGAT2α activity. Representative diseases that may be treated according to the subject methods include: hyperlipidemia (including excess serum triglyceride levels), cardiovascular disease, obesity, diabetes, cancer, neurological disorders, immunological disorders, skin disorders associated with sebaceous gland activity, e.g. acne, and the like.

By treatment is meant at least an amelioration of the symptoms associated with the pathological condition afflicting the host, where amelioration is used in a broad sense to refer to at least a reduction in the magnitude of a parameter, *e.g.* symptom, associated with the pathological condition being treated, such as serum triglyceride level, weight, total body fat content, etc. As such, treatment also includes situations where the pathological condition, or at least symptoms associated therewith, are completely inhibited, *e.g.* prevented from happening, or stopped, *e.g.* terminated, such that the host no longer suffers from the pathological condition, or at least the symptoms that characterize the pathological condition. For example, where the disease condition is marked by the presence of elevated lipid levels, treatment

includes at least a reduction in the observed lipid levels, including a restoration of normal lipid levels. As another example, where the disease is obesity, treatment results in at least a reduction in the overall weight and/or total body fat content of the host.

The subject methods also find use in the modulation of DGAT2α or MGAT1 activity in hosts not suffering from a disease condition but in which the modulation of DGAT2α or MGAT1 activity is nonetheless desired. For example, sperm production in males has been associated with diglyceride acyltransferase activity. As such, in males where at least reduced sperm production is desired, the subject methods can be used to reduce this target activity in such males, e.g. by administering an agent that reduces DGAT2α activity in such males, where such agents are described above. In other words, the subject methods provide a means of male contraception. Alternatively, where increased sperm count in a given male is desired, e.g. in those conditions where the male has reduced fertility, the subject methods can be used to enhance this target activity in the male and thereby increase sperm count and fertility, e.g. by administering to the male host a DGAT2α enhancing agent, as described above.

A variety of hosts are treatable according to the subject methods. Generally such hosts are "mammals" or "mammalian," where these terms are used broadly to describe organisms which are within the class mammalia, including the orders carnivore (*e.g.*, dogs and cats), rodentia (*e.g.*, mice, guinea pigs, and rats), and primates (*e.g.*, humans, chimpanzees, and monkeys). In many embodiments, the hosts will be humans.

Kits with unit doses of the active agent, usually in oral or injectable doses, are provided. In such kits, in addition to the containers containing the unit doses will be an informational package insert describing the use and attendant benefits of the drugs in treating pathological condition of interest. Preferred compounds and unit doses are those described herein above.

METHODS AND COMPOSITIONS FOR PRODUCING DIGLYCERIDES, DIGLYCERIDE COMPOSITIONS, TRIGLYCERIDES, AND TRIGLYCERIDE COMPOSITIONS

Also provided by the subject invention are methods for preparing diglycerides, diglyceride compositions, triglycerides, and triglyceride comprising compositions, as well as

the compositions produced by these methods.

In preparing triglycerides with the subject invention, at least the direct substrates of the desired triacylglyercol, e.g. diacylglycerol and fatty acyl CoA, will be combined in the presence of the polypeptide under conditions sufficient for the acylation of the diacylglycerol to occur. The synthesis may occur in an *in vitro* system, e.g. in a vessel in which the substrates or precursors thereof and the DGAT2α enzyme, as well as any other requisite enzymes (e.g. as need to convert the substrate precursors to substrates), or an *in vivo* system, e.g. a host cell that naturally comprises the substrates and into which a DGAT2α gene or nucleic acid according to the subject invention has been inserted in a manner sufficient for expression of the gene and provision of the DGAT2α enzyme, where the resultant triglyceride products may be separated from the host cell using standard separation techniques.

In preparing diglycerides with the subject invention, at least the direct substrates of the desired diacylglyercol, e.g. monoacylglycerol and fatty acyl CoA, will be combined in the presence of the polypeptide under conditions sufficient for the acylation of the monoacylglycerol to occur. The synthesis may occur in an *in vitro* system, e.g. in a vessel in which the substrates or precursors thereof and the MGAT1 enzyme, as well as any other requisite enzymes (e.g. as need to convert the substrate precursors to substrates), or an *in vivo* system, e.g. a host cell that naturally comprises the substrates and into which an MGAT1 gene or nucleic acid according to the subject invention has been inserted in a manner sufficient for expression of the gene and provision of the MGAT1 enzyme, where the resultant diglyceride products may be separated from the host cell using standard separation techniques.

Of interest for use in producing di- and triglyceride compositions are transgenic plants/fungi that have been genetically manipulated using the nucleic acid compositions of the subject invention to produce di- and/or triglycerides and/or compositions thereof in one or more desirable ways. Transgenic plants/fungi of the subject invention are those plants/fungi that at least: (a) produce more diglyceride, diglyceride composition, triglyceride or triglyceride composition than wild type, e.g. produce more oil, such as by producing seeds having a higher oil content, as compared to wild-type; (b) produce di- or triglyceride compositions, e.g. oils, that are enriched for di- or triglycerides and/or enriched for one or more particular di- or triglycerides as compared to wild type; and the like. Of interest are

transgenic plants that produce commercially valuable triglyceride compositions or oils, such as canola, rapeseed, palm, corn, etc., containing various poly- and mono-unsaturated fatty acids, and the like. Of particular interest are transgenic plants, such as canola, rapeseed, palm, oil, etc., which have been genetically modified to produce seeds having higher oil content than the content found in the corresponding wild type, where the oil content of the seeds produced by such plants is at least 10% higher, usually at least 20 % higher, and in many embodiments at least 30 % higher than that found in the wild type, where in many embodiments seeds having oil contents that are 50% higher, or even greater, as compared to seeds produced by the corresponding wild-type plant, are produced. The seeds produced by such DGAT2α transgenic plants can be used as sources of oil or as sources of additional DGAT2α transgenic plants. Such transgenic plants and seeds therefore find use in methods of producing oils. In such methods, DGAT2α transgenic plants engineered to produce seeds having a higher oil content than the corresponding wild-type, e.g. seeds in which the DGAT2α gene is overexpressed, are grown, the seeds are harvested and then processed to recover the oil. The subject transgenic plants can also be used to produce novel oils characterized by the presence of triglycerides in different amounts and/or ratios than those observed in naturally occurring oils. The transgenic plants/fungi described above can be readily produced by those of skill in the art armed with the nucleic acid compositions of the subject invention. *See* the discussion on how to prepare transgenic plants, *supra*.

The triglyceride compositions described above find use in a variety of different applications. For example, such compositions or oils find use as food stuffs, being used as ingredients, spreads, cooking materials, etc. Alternatively, such oils find use as industrial feedstocks for use in the production of chemicals, lubricants, surfactants and the like.

Also of interest are transgenic non-human animals suitable for use as sources of food products and/or animal based industrial products. Such transgenic non-human animals, e.g. transgenic mice, rats, livestock, such as cows, pigs, horses, birds, etc, may be produced using methods known in the art and reviewed *supra*. Such transgenic non-human animals can be used for sources of a variety of different food and industrial products in which the triglyceride content is specifically tailored in a desirable manner. For example, such transgenic animals that have been modified in a manner such that DGAT2α activity is reduced as compared to

the wild type can be used as sources of food products that are low in triglyceride content, e.g. low fat or lean meat products, low fat milk, low fat eggs, and the like.

The following examples are offered primarily for purposes of illustration. It will be readily apparent to those skilled in the art that the formulations, dosages, methods of administration, and other parameters of this invention may be further modified or substituted in various ways without departing from the spirit and scope of the invention.

EXPERIMENTAL

Example 1: Existence of DGAT2α

A. Mice (DGAT1-/-) lacking DGAT, as described in WO 99/67268 are lean and resistant to diet-induced obesity, but are still capable of synthesizing triglycerides (TG) and have normal plasma TG levels. However, DGAT activity is virtually absent in membrane preparations from DGAT1-/- tissues (Smith et al., Nat.Genet. 2000 (25), 87-90). Using pulse assays in living cells, we measured that the residual TG synthesis activity in DGAT1-/- Mouse Embryonic Fibroblasts (MEF) or adipocytes was about 40% of that in wild-type cells. The results are graphically depicted in Figs. 1A and 1B. In Fig. 1A the membrane fraction isolated from MEF or adipocytes of wild-type or DGAT1-/- mice was used as the enzyme source in DGAT assays *in vitro*. In Fig. 1B living cells were pulse-labeled with [^{14}C]oleic acid for 24 hours and [^{14}C] incorporation in the TG fraction was measured.

In further assays, increased DGAT activity was observed in DGAT1$^{-/-}$ membranes assayed without magnesium; and DGAT activity was observed to vary with magnesium concentration in liver and adipose tissue.

The above findings indicate the existence of DGAT2α, a second enzyme with diglyceride acyltransferase activity.

II. Mammalian DGAT2α Sequences

A. The human DGAT2α nucleic acid and amino acid sequences were identified using standard procedures, as described above. The human DGAT2α cDNA has the sequence appearing as SEQ ID NO:01, infra, while the protein encoded thereby has the sequence appearing as SEQ ID NO:02, infra.

B. The mouse DGAT2α nucleic acid and amino acid sequences were identified using standard procedures, as described above. The mouse DGAT2α cDNA has the sequence appearing as SEQ ID NO:03, infra, while the protein encoded thereby has the sequence appearing as SEQ ID NO:04, infra.

Example 2: Characterization of DGAT2α

A. Molecular Weight

The mouse DGAT2α cDNA was determined to encode a 43 kD predicted protein based on the amino acid sequence. The mouse DGAT2α cDNA was determined to have no sequence homology to DGAT1, as described in Cases et al., supra. The mouse DGAT2α amino acid sequence was determined to have 2 putative N-linked glycosylation sites. The mouse DGAT2α amino acid sequence was determined to have 6 putative PKC phosphorylation sites.A Hydrophobicity plot assessed by Kyte-Doolittle (K-D) analysis revealed the existence of multiple putative transmembrane domains in the mouse DGAT2α amino acid sequence. Fig. 2 provides a graphical result of this analysis. As such, there are regions of higher hydrophobicity compatible with the existence of one or more transmembrane domain.

Example 3: Expression of DGAT2a in insect cells

Sf9 insect cells were infected with wild-type baculovirus, mouse FLAG-tagged DGAT2α or mouse FLAG-tagged DGAT1 (Cases et al., supra) recombinant baculoviruses, and the membrane fractions were assayed for DGAT activity. The results are graphically

provided in Fig. 3A. In Fig. 3A a time course of DGAT2α virus infection is provided. Insect cell membranes were isolated at the indicated times after infection. Expression of the FLAG-tagged DGAT2α protein was detected by immunoblotting with an anti-FLAG antibody (*Inset*). DGAT activity was measured at low (5 mM) or high (100 mM) magnesium concentration, using [^{14}C]oleoyl CoA and cold diacylglycerol. The experiment was repeated three times and a representative experiment is shown. Fig. 3B shows that DGAT2α activity is dependent on the presence of the diacylglycerol substrate. Assays were performed at low magnesium concentration, using [^{14}C]oleoyl CoA with or without exogenous cold diacylglycerol. When no diacylglycerol is added, no significant DGAT activity can be detected over background. Data represent the mean (± SD) of three experiments. To compare the DGAT activity of DGAT1 and DGAT2α , membranes expressing equal levels of DGAT1 or DGAT2α (as assessed by immunoblotting with an anti-FLAG antibody) were assayed at low magnesium concentration using increased amounts of cold oleoyl CoA in the presence of exogenous diacylglycerol. The results are provided in Fig. 3C. Lipids were extracted and separated by TLC and TG accumulation was visualized by charring and quantified by densitometry.

Example 4: Analysis of DGAT2α mRNA expression in various tissues and in adipocyte differentiation

The tissue distribution of human DGAT2α mRNA was analyzed. The results are provided in Fig. 4.

DGAT2α expression increases during 3T3-L1 adipocyte differentiation. Mouse 3T3-L1 adipocyte differentiation was induced and mRNA were isolated at the indicated times shown in Fig. 5. Quantitation of DGAT2α mRNA levels in triplicate samples was performed by Phosphorimager analysis and corrected for loading relative to actin expression. The results are shown in Fig. 5.

Summary of DGAT2α

- mouse DGAT2α has no sequence homology to DGAT1

- mouse DGAT2α diacylglycerol acyltransferase activity inhibited by high magnesium concentrations;
- human DGAT2α RNA expression in many tissues, highest levels found in liver, adipose tissue, and mammary gland
- mouse DGAT2α markedly increased mRNA expression during 3T3-L1 adipocyte differentiation.

Example 5: Characterization of MGAT1 enzymatic activity

The two major pathways for synthesizing diacylglycerol are shown in Figure 9. Figure 9 depicts the roles of DGAT and monoacylglycerol acyltransferases (MGAT) in the synthesis of triacylglycerides. The examples above describe the activity of DGAT2α. In the following section, the activity of MGAT1 is described.

MATERIALS AND METHODS

Cloning of MGAT1 cDNA. Mouse expressed sequence tags (ESTs) encoding MGAT1 were identified by BLAST database searches through their sequence homology to DGAT2 from *Mortierella rammaniana* (accession no. AF391089). Based on these ESTs, primers were designed to amplify the complete coding sequence of MGAT1 from mouse liver RNA by reverse transcription (SuperScript Choice System, Gibco BRL, Rockville, Maryland) and PCR (Takara *Ex Taq*, Panvera, Madison, Wisconsin). The MGAT1 sequence has been deposited in GenBank (accession no. AF384162).

Insect Cell Expression Studies. MGAT1 was tagged with an N-terminal FLAG epitope (MGDYKDDDDG, epitope underlined; SEQ ID NO:17) and expressed in *Spodoptera frugiperda* Sf9 insect cells as described. MGAT1 without FLAG was also expressed to determine whether the presence of FLAG, which permits the detection and assessment of expression levels, alters MGAT1 activity. Briefly, the MGAT1 coding sequence (with or without FLAG) was subcloned into pVL1393 baculovirus transfer vector (PharMingen, San Diego, CA). Recombinant baculoviruses were generated by cotransfecting Sf9 insect cells with the transfer vector and BaculoGold DNA (PharMingen). High-titer viruses used for

MGAT1 expression were obtained after two rounds of amplification. FLAG-tagged-DGAT1 (accession no. AF078752) and -DGAT2 (accession no. AF384160) were also expressed as controls. To prepare membrane fractions, cells were typically infected with virus for 3 days, washed with PBS, and homogenized by 10 passages through a 27-gauge needle in 1 mM EDTA, 200 mM sucrose, 100 mM Tris-HCl, pH 7.4. Total membrane fractions (100,000 x *g* pellet) were resuspended in homogenization buffer and frozen at –80°C until use. Expression of MGAT1, DGAT2, and FLAG-tagged proteins in 5 μg of membrane protein was verified by immunoblotting with an antiserum raised against the C-terminus (amino acids 295–316) of MGAT1, an antiserum against DGAT2, and an anti-Flag M2 antibody (Sigma, St. Louis, Missouri), respectively.

***In Vitro* Acyltransferase Assays.** Generally, acyltransferase activities were assayed under apparent V_{max} conditions for 5 min in a final volume of 200 μl. Each reaction contained 100 μg of membrane proteins, 5 mM $MgCl_2$, 1.25 mg/ml bovine serum albumin (BSA), 200 mM sucrose, 100 mM Tris-HCl, pH 7.4, 25 μM acyl donor, and 200 μM acyl acceptor. Non-polar acyl acceptors (diacylglycerol, monoacylglycerol, cholesterol, phosphatidic acid, and sphingosine) were dispersed as phosphatidylcholine liposomes (molar ratio ≈ 0.2), and polar acyl acceptors (glycerol-3-phosphate, dihydroxyacetone phosphate, lysophosphatidic acid, and lysophosphatidylcholine) were dissolved in water. Reactions were started by adding protein and terminated by adding 4 ml of chloroform:methanol (2:1 v/v). The extracted lipids were dried, separated by TLC with hexane:ethyl ether:acetic acid (80:20:1 v/v/v), visualized with iodine vapor, and identified with lipid standards. For experiments with radiolabeled substrates, TLC plates were exposed to x-ray film to assess the incorporation of radioactivity into lipid products.

Specifically, DGAT activity was measured as described. Cases et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:13018-13023. MGAT activity was determined by measuring the incorporation of the [^{14}C]oleoyl moiety into diacylglycerol with 25 μM [^{14}C]oleoyl CoA (specific activity, ~20,000 dpm/nmol) and 200 μM exogenously added *sn*-2 monooleoylglycerol. In some assays, [^{14}C]*sn*-1 monooleoylglycerol (specific activity, 18 μCi/μmol, 200 μM final concentration, American Radiolabeled Chemicals, St. Louis,

Missouri) was used as a radiolabeled tracer to measure MGAT activity in the presence of unlabeled oleoyl CoA (25 μM). The dependence of MGAT1 activity on monoacylglycerol and fatty acyl CoA as substrates was determined by assaying MGAT1 with various concentrations of oleoyl CoA or monooleoylglycerol in the presence of 400μM monooleoylglycerol or 50 μM oleoyl CoA, respectively. Diacylglycerol mass was quantified by densitometry after the lipid products were separated by TLC and visualized by immersing the plate in a solution of 10% cupric sulfate and 8% phosphoric acid and heating at 180°C for 30 min. Stereoisomers of monoacylglycerol (*sn*-1-, *sn*-2-monooleoylglycerol, and 3-monostearoylglycerol) were from Sigma. MGAT activity in tissues was measured in particulate fractions prepared from pooled tissues of three 15-week-old male mice.

Mammalian Cell Expression Studies. For mammalian cell expression, FLAG-tagged MGAT1 was subcloned into a pcDNA3 vector and transfected into COS-7 or CHO cells with Fugene 6 (Roche Diagnostics, Chicago, Illinois). FLAG-tagged-DGAT1, -DGAT2, and -ACAT2 (cholesterol acyltransferase, accession no. AF078751) were expressed as controls. Membrane fractions were prepared as described for insect cells. Expression of FLAG-tagged proteins (in 20 μg of membrane proteins) was verified by immunoblotting with the anti-FLAG M2 antibody. For immunocytochemistry, cells were grown and transfected on glass cover slips. Two days after transfection, cells were fixed in acetone:methanol (1:1) for 2 min and incubated in PBS containing 3% BSA and 0.2% Triton X-100 for 1 h at room temperature. Samples were then incubated sequentially with 4 μg/ml anti-FLAG antibody (Sigma) for 1 h and 10 μg/ml FITC-conjugated goat anti-mouse IgG (CalBiochem, Pasadena, California) for 30 min. Antibodies were diluted in PBS containing 3% BSA and 0.02% Triton X-100. MGAT activities in membranes of transfected cells were assayed as described above.

MGAT1 Tissue Expression Pattern in Mice. To determine tissue distribution of MGAT1 expression, a mouse multiple tissue blot (SeeGene, Seoul, Korea), a blot of total RNA from indicated tissues, and a poly-A^+ RNA blot (CLONTECH, Palo Alto, California) were hybridized with ^{32}P-labeled probes generated by random priming (Amersham) with MGAT1 cDNA as the template.

Results

Identification of an MGAT1 Gene. Mouse MGAT1 cDNA was originally identified as a DGAT candidate (DC) gene through its homology to genes encoding DGAT2. Mouse MGAT1 is also referred to herein, and in Cases et al. ((2001) *J. Biol. Chem.* 276:38870-38876), as mouse DC2. The amino acid sequence of mouse MGAT1 is set forth in SEQ ID NO:06; and the nucleotide sequence encoding mouse MGAT1 is set forth in SEQ ID NO:05. The open reading frame of the MGAT1 cDNA predicts a 335–amino acid protein, which is 40% identical to mouse DGAT2, as shown in Figure 10A, and has a predicted molecular mass of 38.8 kDa. Like DGAT2, MGAT1 contains sequences similar to a domain of phosphate acyltransferases. MGAT1 also possesses two putative N-linked glycosylation sites and a potential tyrosine phosphorylation site. The hydrophobicity plot for MGAT1 is similar to that for DGAT2 and predicts at least one transmembrane domain (amino acids 21–43) in the amino terminus, as shown in Figure 10B. Sequences for the human MGAT1 homologue have been reported as human DC2, a member of the DGAT2 gene family. The mouse MGAT1 gene is located on chromosome 1 (accession no. AC079223) and its human homologue is on chromosome 2 (accession no. NT_005126).

Mouse MGAT1 expressed in Insect Cells. To examine the biochemical activity of MGAT1 protein, we expressed FLAG epitope–tagged and non-tagged versions of the cDNA in insect cells. The non-FLAG-tagged version migrated on SDS-PAGE with an apparent molecular mass of ~33 kDa. As expected, the FLAG-tagged version migrated slightly more slowly because of the FLAG epitope. Because MGAT1 shares sequence homology with DGAT2, we first examined whether membranes expressing MGAT1 have DGAT activity. With either [^{14}C]dioleoylglycerol or [^{14}C]oleoyl CoA as a radiolabeled substrate, MGAT1-expressing membranes incorporated more radioactivity into triacylglycerols than membranes expressing wild-type viral proteins or heat-inactivated MGAT1 indicating that these membranes have DGAT activity. However, this DGAT activity was significantly less than that in control membranes expressing DGAT2, even though MGAT1 protein was expressed at a higher level.

Since the assays with [^{14}C]oleoyl CoA radiolabel revealed that a significant amount of [^{14}C]oleoyl CoA was incorporated into diacylglycerol in membranes expressing MGAT1, we

suspected that MGAT1 possesses MGAT activity. To test this possibility, membranes expressing MGAT1 were assayed with either [^{14}C]monooleoylglycerol or [^{14}C]oleoyl CoA as the labeled substrate. In both cases, membranes expressing MGAT1 catalyzed the incorporation of the label into diacylglycerol, establishing that the MGAT1 protein possesses MGAT activity. This MGAT activity was confirmed by its dependence on MGAT substrates; when unlabeled MGAT substrate (monooleoylglycerol or oleoyl CoA) was provided over a range of concentrations while the other substrate was held constant, the mass of diacylglycerol synthesized was dependent on the concentration of substrate, as shown in Figure 11. Further, the acyltransferase activity of MGAT1 appeared to be specific for monoacylglycerol (and possibly diacylglycerol) as the acyl group acceptor; no acyltransferase activity was found in MGAT1-expressing membranes when glycerol-3-phosphate, dihydroxyacetone phosphate, lysophosphatidate, lysophosphatidylcholine, sphingosine, or cholesterol were used as the [^{14}C]oleoyl CoA acceptor.

Next, since there are three stereoisomers of monoacylglycerol, we determined whether MGAT1 can acylate each of these stereoisomers. In these assays [^{14}C]oleoyl CoA was used as the acyl donor and either *sn*-1-monooleoylglycerol, *sn*-2-monooleoylglycerol, or *sn*-3-monostearoylglycerol as the acyl acceptor. When *sn*-1-monooleoylglycerol or *sn*-3-monostearoylglycerol was used, the major product was *sn*-1,3- diacylglycerol. On the other hand, when *sn*-2-monooleoylglycerol was used, the major product was *sn*-1,2(2,3)-diacylglycerol. These results indicate that MGAT1 can acylate each of the stereoisomers of monoacylglycerol, mainly at the *sn*-1 or *sn*-3 position. We also found that the specific activities of MGAT1 using either *sn*-1-monooleoylglycerol or *sn*-2-monooleoylglycerol as substrates were similar and that both increased proportionally with MGAT1 protein levels, as shown in Figure 12. These findings indicate that MGAT1 expressed in insect cells can use *sn*-1-monooleoylglycerol and *sn*-2-monooleoylglycerol equally well as substrates *in vitro*.

Mouse MGAT1 expressed in Mammalian Cells. To examine whether MGAT1 protein expressed in mammalian cells also has MGAT activity, we expressed mouse MGAT1 cDNA and control cDNAs transiently in monkey kidney COS-7 cells, as shown in Figure 13. In cells expressing MGAT1, immunofluorescence microscopy demonstrated a perinuclear and

reticular staining pattern, consistent with distribution of the protein in the endoplasmic reticulum. COS-7 cell membranes expressing MGAT1 incorporated more radioactivity into diacylglycerol when [^{14}C]oleoyl CoA and *sn*-2-monooleoylglycerol were provided as substrates, indicating that these membranes possess MGAT activity. Similar levels of MGAT activity were found when *sn*-1-monooleoylglycerol was used as the acyl acceptor. In contrast, MGAT activity was not detected in control membranes, except in those expressing DGAT1, which appeared to possess a low level of MGAT activity. Similar results were found when these cDNAs were expressed in Chinese hamster ovary (CHO) cells.

Tissue Distribution of Mouse MGAT1 Expression. MGAT1 mRNA expression was highest in the stomach and kidney; lower levels of expression were present in white and brown adipose tissue, uterus, and liver. MGAT1 mRNA expression was not detected in the small intestine. Because MGAT activity had not previously been reported in mice, we performed MGAT assays on membranes from mouse tissues. As demonstrated for many other species, the highest activity was found in the small intestine, as shown in Figure 14. MGAT activity was also detected at significant levels in the stomach, kidney, adipose tissue, and liver, where MGAT1 is expressed.

It is apparent from the above results and discussion that polynucleotides encoding mammalian DGAT2α and MGAT1 enzymes, as well as novel polypeptides encoded thereby, are provided. The subject invention is important for both research and therapeutic applications. Using the DGAT2α probes of the subject invention, the role of DGAT2α and its regulation in a number of physiological processes can be studied *in vivo*. The subject invention also provides for important new ways of treating diseases associated with DGAT2α and MGAT, such as hypertriglycemia and obesity, as well as in the production of tryglycerides.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an

admission that the present invention is not entitled to antedate such publication by virtue of prior invention.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.